



OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/966,803

DATE: 03/12/2002

TIME: 09:44:20

Input Set : N:\Crf3\RULE60\09966803.raw

Output Set: N:\CRF3\03122002\I966803.raw

SEQUENCE LISTING

1 (1) GENERAL INFORMATION:

2 (i) APPLICANT: Murphy et al.

3 (ii) TITLE OF INVENTION: Amidases

4 (iii) NUMBER OF SEQUENCES: 4

5 (iv) CORRESPONDENCE ADDRESS:

6 (A) ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
7 CECCHI, STEWART & OLSTEIN

8 (B) STREET: 6 BECKER FARM ROAD

9 (C) CITY: ROSELAND

10 (D) STATE: NEW JERSEY

11 (E) COUNTRY: USA

12 (F) ZIP: 07068

13 (v) COMPUTER READABLE FORM:

14 (A) MEDIUM TYPE: 3.5 INCH DISKETTE

15 (B) COMPUTER: IBM PS/2

16 (C) OPERATING SYSTEM: MS-DOS

17 (D) SOFTWARE: WORD PERFECT 5.1

18 (vi) CURRENT APPLICATION DATA:

C--> 19 (A) APPLICATION NUMBER: US/09/966,803

C--> 20 (B) FILING DATE: 27-Sep-2001

21 (C) CLASSIFICATION:

22 (vii) PRIOR APPLICATION DATA:

23 (A) APPLICATION NUMBER: US/09/609,570

24 (B) FILING DATE: 30-Jun-2000

25 (A) APPLICATION NUMBER: 09/427,372

26 (B) FILING DATE:

27 (viii) ATTORNEY/AGENT INFORMATION:

28 (A) NAME: Charles J. Herron

29 (B) REGISTRATION NUMBER: 28,019

30 (C) REFERENCE/DOCKET NUMBER: 331400-53

31 (ix) TELECOMMUNICATION INFORMATION:

32 (A) TELEPHONE: 201-994-1700

33 (B) TELEFAX: 201-994-1744

34 (2) INFORMATION FOR SEQ ID NO: 1:

35 (i) SEQUENCE CHARACTERISTICS:

36 (A) LENGTH: 1869 NUCLEOTIDES

37 (B) TYPE: NUCLEIC ACID

38 (C) STRANDEDNESS: SINGLE

39 (D) TOPOLOGY: LINEAR

W--> 40 (ii) MOLECULE TYPE: DNA

41 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

42 ATG ACC GGC ATC GAA TGG AAC CAC GAG ACC TTT TCT AAG TTC GCC TAC 48

ENTERED

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/966,803

DATE: 03/12/2002

TIME: 09:44:20

Input Set : N:\Crf3\RULE60\09966803.raw

Output Set: N:\CRF3\03122002\I966803.raw

```

W--> 43      Met Thr Tly Ile Glu Trp Asn His Glu Thr Phe Ser Lys Phe Ala Tyr
44              5              10              15
45      CTG GGC GAC CCG AGG ATA CGG GGA AAC TTA ATC GCG TAC ACC CTG ACG 96
46      Leu Gly Asp Pro Arg Ile Arg Gly Asn Leu Ile Ala Tyr Thr Leu Thr
47              20              25              30
48      AAG GCC AAC ATG AAG GAC AAC AAG TAC GAG AGC ACG GTT GTT GTT GAA 144
49      Lys Ala Asn Met Lys Asp Asn Lys Tyr Glu Ser Thr Val Val Val Glu
50              35              40              45
51      GAC CTT GAA ACG GGC TCA AGG CGC TTC ATC GAG AAC GCC TCA ATG CCG 192
52      Asp Leu Glu Thr Gly Ser Arg Arg Phe Ile Glu Asn Ala Ser Met Pro
53              50              55              60
54      AGG ATT TCG CCA GAC GGC AGA AAG CTC GCC TTC ACC TGC TTT AAC GAG 240
55      Arg Ile Ser Pro Asp Gly Arg Lys Leu Ala Phe Thr Cys Phe Asn Glu
56              65              70              75              80
57      GAG AAG AAG GAG ACC GAG ATA TGG GTG GCC GAT ATC CAG ACC CTG AGC 288
58      Glu Lys Lys Glu Thr Glu Ile Trp Val Ala Asp Ile Gln Thr Leu Ser
59              85              90              95
60      GCC AAG AAA GTC CTC TCA ACT AAA AAC GTC CGC TCG ATG CAG TGG AAC 336
61      Ala Lys Lys Val Leu Ser Thr Lys Asn Val Arg Ser Met Gln Trp Asn
62              100              105              110
63      GAC GAT TCA AGG AGA CTC TTA GTT GTC GGC TTC AAG AGG AGG GAC GAT 384
64      Asp Asp Ser Arg Arg Leu Leu Val Val Gly Phe Lys Arg Arg Asp Asp
65              115              120              125
66      GAG GAC TTC GTC TTT GAC GAC GAC GTC CCG GTC TGG TTC GAC AAT ATG 432
67      Glu Asp Phe Val Phe Asp Asp Asp Val Pro Val Trp Phe Asp Asn Met
68              130              135              140
69      GGA TTC TTT GAT GGA GAG AAG ACG ACG TTC TGG GTT CTT GAC ACT GAG 480
70      Gly Phe Phe Asp Gly Glu Lys Thr Thr Phe Trp Val Leu Asp Thr Glu
71      145              150              155              160
72      GCC GAG GAG ATA ATC GAG CAG TTC GAG AAG CCG AGG TTT TCG AGT GGC 528
73      Ala Glu Glu Ile Ile Glu Gln Phe Glu Lys Pro Arg Phe Ser Ser Gly
74              165              170              175
75      CTC TGG CAC GGC GAT GCG ATA GTT GTG AAC GTC CCG CAC CGC GAG GGG 576
76      Leu Trp His Gly Asp Ala Ile Val Val Asn Val Pro His Arg Glu Gly
77              180              185              190
78      AGC AAG CCT GCC CTG TTC AAG TTC TAC GAC ATA GTC CTA TGG AAG GAC 624
79      Ser Lys Pro Ala Leu Phe Lys Phe Tyr Asp Ile Val Leu Trp Lys Asp
80              195              200              205
81      GGG GAG GAA GAG AAG CTC TTC GAG AGG GTC TCC TTC GAG GCG GTT GAC 672
82      Gly Glu Glu Glu Lys Leu Phe Glu Arg Val Ser Phe Glu Ala Val Asp
83              210              215              220
84      TCC GAC GGA AAG AGA ATA CTC CTG AGG GGC AAG AAA AAA AAG CGG TTC 720
85      Ser Asp Gly Lys Arg Ile Leu Leu Arg Gly Lys Lys Lys Lys Arg Phe
86      225              230              235              240
87      ATC AGC GAG CAC GAC TGG CTG TAC CTC TGG GAC GGC GAG CTT AAA CCG 768
88      Ile Ser Glu His Asp Trp Leu Tyr Leu Trp Asp Gly Glu Leu Lys Pro
89              245              250              255
90      ATC TAC GAG GGC CCG CTC GAC GTC TGG GAA GCC AAG CTC ACG GAA GGA 816
91      Ile Tyr Glu Gly Pro Leu Asp Val Trp Glu Ala Lys Leu Thr Glu Gly

```

## RAW SEQUENCE LISTING

DATE: 03/12/2002

PATENT APPLICATION: US/09/966,803

TIME: 09:44:20

Input Set : N:\Crf3\RULE60\09966803.raw

Output Set: N:\CRF3\03122002\I966803.raw

```

92          260          265          270
93 AAG GTC TAC TTC CTC ACT CCA GAT GCG GGC AGG GTA AAC CTC TGG CTC 864
94 Lys Val Tyr Phe Leu Thr Pro Asp Ala Gly Arg Val Asn Leu Trp Leu
95          275          280          285
96 TGG GAC GGG AAG GCC GAG CGT GTT GTT ACC GGC GAC CAC TGG ATT TAC 912
97 Trp Asp Gly Lys Ala Glu Arg Val Val Thr Gly Asp His Trp Ile Tyr
98          290          295          300
99 GGG CTT GAC GTC AGC GAT GGC AAA GCA TTG CTC CTC ATC ATG ACC GCC 960
100 Gly Leu Asp Val Ser Asp Gly Lys Ala Leu Leu Leu Ile Met Thr Ala
101          305          310          315          320
102 ACG AGG ATA GGC GAG CTC TAC CTC TAC GAC GGC GAG CTG AAA CAG GTC 1008
103 Thr Arg Ile Gly Glu Leu Tyr Leu Tyr Asp Gly Glu Leu Lys Gln Val
104          325          330          335
105 ACC GAA TAC AAC GGC CCG ATA TTC AGG AAG CTC AAG ACC TTC GAG CCG 1056
106 Thr Glu Tyr Asn Gly Pro Ile Phe Arg Lys Leu Lys Thr Phe Glu Pro
107          340          345          350
108 AGG CAC TTC CGC TTC AAG AGC AAA GAC CTC GAG ATA GAC GGC TGG TAC 1104
109 Arg His Phe Arg Phe Lys Ser Lys Asp Leu Glu Ile Asp Gly Trp Tyr
110          355          360          365
111 CTC AGG CCG GAG GTT AAA GAG GAG AAG GCC CCG GTG ATA GTC TTC GTC 1152
112 Leu Arg Pro Glu Val Lys Glu Glu Lys Ala Pro Val Ile Val Phe Val
113          370          375          380
114 CAC GGC GGG CCG AAG GGC ATG TAC GGA CAC CGC TTC GTC TAC GAG ATG 1200
115 His Gly Gly Pro Lys Gly Met Tyr Gly His Arg Phe Val Tyr Glu Met
116          385          390          395          400
117 CAG CTG ATG GCG AGC AAG GGC TAC TAC TGC TGC TTC GTG AAC CCG CGC 1248
118 Gln Leu Met Ala Ser Lys Gly Tyr Tyr Val Val Phe Val Asn Pro Arg
119          405          410          415
120 GGC AGC GAC GGC TAT AGC GAA GAC TTC GCG CTC CGC GTC CTG GAG AGG 1296
121 Gly Ser Asp Gly Tyr Ser Glu Asp Phe Ala Leu Arg Val Leu Glu Arg
122          420          425          430
123 ACT GGC TTG GAG GAC TTT GAG GAC ATA ATG AAC GGC ATC GAG GAG TTC 1344
124 Thr Gly Leu Glu Asp Phe Glu Asp Ile Met Asn Gly Ile Glu Glu Phe
125          435          440          445
126 TTC AAG CTC GAA CCG CAG GCC GAC AGG GAG CGC GTT GGA ATA ACG GGC 1392
127 Phe Lys Leu Glu Pro Gln Ala Asp Arg Glu Arg Val Gly Ile Thr Gly
128          450          455          460
129 ATA AGC TAC GGC GGC TTC ATG ACC AAC TGG GCC TTG ACT CAG AGC GAC 1440
130 Ile Ser Tyr Gly Gly Phe Met Thr Asn Trp Ala Leu Thr Gln Ser Asp
131          465          470          475          480
132 CTC TTC AAG GCA GGA ATA AGC GAG AAC GGC ATA AGC TAC TGG CTC ACC 1488
133 Leu Phe Lys Ala Gly Ile Ser Glu Asn Gly Ile Ser Tyr Trp Leu Thr
134          485          490          495
135 AGC TAC GCC TTC TCG GAC ATA GGG CTC TGG TAC GAC GTC GAG GTC ATC 1536
136 Ser Tyr Ala Phe Ser Asp Ile Gly Leu Trp Tyr Asp Val Glu Val Ile
137          500          505          510
138 GGG CCA AAT CCG TTA GAG AAC GAG AAC TTC AGG AAG CTC AGC CCG CTG 1584
139 Gly Pro Asn Pro Leu Glu Asn Glu Asn Phe Arg Lys Leu Ser Pro Leu
140          515          520          525

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/966,803

DATE: 03/12/2002

TIME: 09:44:20

Input Set : N:\Crf3\RULE60\09966803.raw

Output Set: N:\CRF3\03122002\I966803.raw

```

141   TTC TAC GCT CAG AAC GTG AAG GCG CCG ATA CTC CTA ATC CAC TCG CTT 1632
142   Phe Tyr Ala Gln Asn Val Lys Ala Pro Ile Leu Leu Ile His Ser Leu
143       530                      535                      540
144   GAG GAC TAC CGC TGT CCG CTC GAC CAG AGC CTT ATG TTC TAC AAC GTG 1680
145   Glu Asp Tyr Arg Cys Pro Leu Asp Gln Ser Leu Met Phe Tyr Asn Val
146       545                      550                      555                      560
147   CTC AAG GAC ATG GGC AAG GAA GCC TAC ATA GCG ATA TTC AAG CGC GGC 1728
148   Leu Lys Asp Met Gly Lys Glu Ala Tyr Ile Ala Ile Phe Lys Arg Gly
149       565                      570                      575
150   GCC CAC GGC CAC AGC GTC CGC GGA AGC CCG AGG CAC AGG CCG AAG CGC 1776
151   Ala His Gly His Ser Val Arg Gly Ser Pro Arg His Arg Pro Lys Arg
152       580                      585                      590
153   TAC AGG CTC TTC ATA GAG TTC TTC GAG CGC AAG CTC AAG AAG TAC GAG 1824
154   Tyr Arg Leu Phe Ile Glu Phe Phe Glu Arg Lys Leu Lys Lys Tyr Glu
155       595                      600                      605
156   GAG GGC TTT GAG GTA GAG AAG ATA CTC AAG GGG AAT GGG AAC TGA      1869
157   Glu Gly Phe Glu Val Glu Lys Ile Leu Lys Gly Asn Gly Asn
158       610                      615                      620

160 (2) INFORMATION FOR SEQ ID NO: 2:
161   (i) SEQUENCE CHARACTERISTICS:
162       (A) LENGTH: 622 AMINO ACIDS
163       (B) TYPE: AMINO ACID
164       (C) STRANDEDNESS:
165       (D) TOPOLOGY: LINEAR
166   (ii) MOLECULE TYPE: PROTEIN
167   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
168   Met Thr Gly Ile Glu Trp Asn His Glu Thr Phe Ser Lys Phe Ala Tyr
169           5                      10                      15
170   Leu Gly Asp Pro Arg Ile Arg Gly Asn Leu Ile Ala Tyr Thr Leu Thr
171           20                      25                      30
172   Lys Ala Asn Met Lys Asp Asn Lys Tyr Glu Ser Thr Val Val Val Glu
173           35                      40                      45
174   Asp Leu Glu Thr Gly Ser Arg Arg Phe Ile Glu Asn Ala Ser Met Pro
175           50                      55                      60
176   Arg Ile Ser Pro Asp Gly Arg Lys Leu Ala Phe Thr Cys Phe Asn Glu
177           65                      70                      75                      80
178   Glu Lys Lys Glu Thr Glu Ile Trp Val Ala Asp Ile Gln Thr Leu Ser
179           85                      90                      95
180   Ala Lys Lys Val Leu Ser Thr Lys Asn Val Arg Ser Met Gln Trp Asn
181           100                     105                     110
182   Asp Asp Ser Arg Arg Leu Leu Val Val Gly Phe Lys Arg Arg Asp Asp
183           115                     120                     125
184   Glu Asp Phe Val Phe Asp Asp Asp Val Pro Val Trp Phe Asp Asn Met
185           130                     135                     140
186   Gly Phe Phe Asp Gly Glu Lys Thr Thr Phe Trp Val Leu Asp Thr Glu
187           145                     150                     155                     160
188   Ala Glu Glu Ile Ile Glu Gln Phe Glu Lys Pro Arg Phe Ser Ser Gly
189           165                     170                     175
190   Leu Trp His Gly Asp Ala Ile Val Val Asn Val Pro His Arg Glu Gly

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/966,803

DATE: 03/12/2002

TIME: 09:44:20

Input Set : N:\Crif3\RULE60\09966803.raw

Output Set: N:\CRF3\03122002\I966803.raw

191		180		185		190
192	Ser Lys Pro Ala Leu Phe Lys Phe Tyr Asp Ile Val Leu Trp Lys Asp					
193		195		200		205
194	Gly Glu Glu Glu Lys Leu Phe Glu Arg Val Ser Phe Glu Ala Val Asp					
195		210		215		220
196	Ser Asp Gly Lys Arg Ile Leu Leu Arg Gly Lys Lys Lys Lys Arg Phe					
197	225		230		235	240
198	Ile Ser Glu His Asp Trp Leu Tyr Leu Trp Asp Gly Glu Leu Lys Pro					
199		245		250		255
200	Ile Tyr Glu Gly Pro Leu Asp Val Trp Glu Ala Lys Leu Thr Glu Gly					
201		260		265		270
202	Lys Val Tyr Phe Leu Thr Pro Asp Ala Gly Arg Val Asn Leu Trp Leu					
203		275		280		285
204	Trp Asp Gly Lys Ala Glu Arg Val Val Thr Gly Asp His Trp Ile Tyr					
205		290		295		300
206	Gly Leu Asp Val Ser Asp Gly Lys Ala Leu Leu Leu Ile Met Thr Ala					
207	305		310		315	320
208	Thr Arg Ile Gly Glu Leu Tyr Leu Tyr Asp Gly Glu Leu Lys Gln Val					
209		325		330		335
210	Thr Glu Tyr Asn Gly Pro Ile Phe Arg Lys Leu Lys Thr Phe Glu Pro					
211		340		345		350
212	Arg His Phe Arg Phe Lys Ser Lys Asp Leu Glu Ile Asp Gly Trp Tyr					
213		355		360		365
214	Leu Arg Pro Glu Val Lys Glu Glu Lys Ala Pro Val Ile Val Phe Val					
215		370		375		380
216	His Gly Gly Pro Lys Gly Met Tyr Gly His Arg Phe Val Tyr Glu Met					
217	385		390		395	400
218	Gln Leu Met Ala Ser Lys Gly Tyr Tyr Val Val Phe Val Asn Pro Arg					
219		405		410		415
220	Gly Ser Asp Gly Tyr Ser Glu Asp Phe Ala Leu Arg Val Leu Glu Arg					
221		420		425		430
222	Thr Gly Leu Glu Asp Phe Glu Asp Ile Met Asn Gly Ile Glu Glu Phe					
223		435		440		445
224	Phe Lys Leu Glu Pro Gln Ala Asp Arg Glu Arg Val Gly Ile Thr Gly					
225		450		455		460
226	Ile Ser Tyr Gly Gly Phe Met Thr Asn Trp Ala Leu Thr Gln Ser Asp					
227	465		470		475	480
228	Leu Phe Lys Ala Gly Ile Ser Glu Asn Gly Ile Ser Tyr Trp Leu Thr					
229		485		490		495
230	Ser Tyr Ala Phe Ser Asp Ile Gly Leu Trp Tyr Asp Val Glu Val Ile					
231		500		505		510
232	Gly Pro Asn Pro Leu Glu Asn Glu Asn Phe Arg Lys Leu Ser Pro Leu					
233		515		520		525
234	Phe Tyr Ala Gln Asn Val Lys Ala Pro Ile Leu Leu Ile His Ser Leu					
235		530		535		540
236	Glu Asp Tyr Arg Cys Pro Leu Asp Gln Ser Leu Met Phe Tyr Asn Val					
237	545		550		555	560
238	Leu Lys Asp Met Gly Lys Glu Ala Tyr Ile Ala Ile Phe Lys Arg Gly					
239		565		570		575

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/966,803

DATE: 03/12/2002

TIME: 09:44:21

Input Set : N:\Crf3\RULE60\09966803.raw

Output Set: N:\CRF3\03122002\I966803.raw

L:19 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]  
L:20 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]  
L:40 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=1  
L:43 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:1  
L:253 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=3  
L:263 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=4